

CALDERÓN COMMUTATORS AND THE CAUCHY INTEGRAL ON LIPSCHITZ CURVES REVISITED III. POLYDISC EXTENSIONS

CAMIL MUSCALU

ABSTRACT. This article is the last in a series of three papers, whose scope is to give new proofs to the well known theorems of Calderón, Coifman, McIntosh and Meyer [1], [6], [7]. Here we extend the results of the previous two papers to the polydisc setting. In particular, we solve completely an open question of Coifman from the early eighties.

1. INTRODUCTION

The present article is a natural continuation of the previous [19], [20] and is the last paper in the sequel. The goal of it is to show that the method developed in these papers to give new proofs to the L^p boundedness of the Calderón commutators and the Cauchy integral on Lipschitz curves [1], [6], [7], can be used to extend these classical results to the n -parameter polydisc setting, for any $n \geq 2$.

Suppose that F is an analytic function on a disc of a certain radius centered at the origin in the complex plane and A a complex valued function in \mathbf{R}^n , so that $\frac{\partial^n A}{\partial x_1 \dots \partial x_n} \in L^\infty(\mathbf{R}^n)$ with an L^∞ norm strictly smaller than the radius of convergence of F . Define the linear operator $C_{n,F,A}$ by the formula

$$C_{n,F,A}f(x) := p.v. \int_{\mathbf{R}^n} f(x+t) F\left(\frac{\Delta_{t_1}^{(1)}}{t_1} \circ \dots \circ \frac{\Delta_{t_n}^{(n)}}{t_n} A(x)\right) \frac{dt_1}{t_1} \dots \frac{dt_n}{t_n} \quad (1)$$

for functions of n variables $f(x)$ for which the principal value integral exists, where $\Delta_s^{(i)}$ denotes the finite difference operator at scale s in the direction of e_i , given by

$$\Delta_s^{(i)} B(x) := B(x + se_i) - B(x)$$

and e_1, \dots, e_n is the standard basis in \mathbf{R}^n .

The main theorem we are going to prove is the following.

Theorem 1.1. *The operator $C_{n,F,A}$ extends naturally as a bounded linear operator from $L^p(\mathbf{R}^n)$ into $L^p(\mathbf{R}^n)$ for every $1 < p < \infty$.*

This answers completely an open question of Coifman from the early eighties [10], [11]. The case when the L^∞ norm of $\frac{\partial^n A}{\partial x_1 \dots \partial x_n}$ is small and the generic $n = 2$ case, have been understood earlier by Journé in [10] and [11] respectively. Our proof is quite different from the approach in [10], [11] and works equally well in all dimensions. In fact, as we will describe in the last section of the paper, much more can be proved in the same way. Not only the operators of (1) are bounded, but also (for instance) those given by expressions of type

$$f \rightarrow p.v. \int_{\mathbf{R}^4} f(x+t+s) F\left(\frac{\Delta_{t_1}^{(1)}}{t_1} \circ \frac{\Delta_{t_2}^{(2)}}{t_2} \circ \frac{\Delta_{s_1}^{(1)}}{s_1} \circ \frac{\Delta_{s_2}^{(2)}}{s_2} A(x)\right) \frac{dt_1}{t_1} \frac{dt_2}{t_2} \frac{ds_1}{s_1} \frac{ds_2}{s_2} \quad (2)$$

and their natural generalizations. Of course, in (2) one has to assume this time that $\frac{\partial^4 A}{\partial x_1^2 \partial x_2^2} \in L^\infty(\mathbf{R}^2)$. When $F(z) = z^d$ with $d \geq 1$ the operator in (1) is the natural n -parameter extension of the d th Calderón commutator, whereas for $F(z) = \frac{1}{1+iz}$ one obtains the n -parameter generalization of the Cauchy integral on Lipschitz curves [1], [6], [7].

For simplicity, we shall denote from now on with $C_{n,d,A}$ the n -parameter d th Calderón commutator. It is easy to observe that when $f(x)$ and $A(x)$ are particularly given by

$$f(x) = f_1(x_1) \cdot \dots \cdot f_n(x_n)$$

and

$$A(x) = A_1(x_1) \cdot \dots \cdot A_n(x_n)$$

one has

$$C_{n,d,A}f(x) = C_{1,d,A_1}f_1(x_1) \cdot \dots \cdot C_{1,d,A_n}f_n(x_n).$$

To motivate the introduction of the operators $C_{n,F,A}$ one just has to recall the context in which the original Calderón commutators appeared [1], [2], [6]. If one tries to extend Calderón's algebra to \mathbf{R}^n and to include in it pseudodifferential operators containing partial derivatives, one is naturally led to the study of the operators in (1) and their natural generalizations.

It is clear and very well known that to prove statements such as the one in Theorem 1.1, one needs to prove *polynomial bounds* for the corresponding Calderón commutators $C_{n,d,A}$. More specifically, Theorem 1.1 reduces to the estimate

$$\|C_{n,d,A}f\|_p \leq C(n,d) \cdot C(p) \cdot \|f\|_p \cdot \left\| \frac{\partial^n A}{\partial x_1 \dots \partial x_n} \right\|_\infty^d \quad (3)$$

for any $f \in L^p$, where $C(n,d)$ grows at most polynomially in d ¹.

The argument of [10] to prove the *small L^∞ norm* theorem used an induction on the dimension n . We work instead directly in \mathbf{R}^n and since our method is essentially similar in every dimension, to keep the technicalities to a minimum, we chose for the reader's convenience to describe the proof of the main Theorem 1.1 in the particular case of the plane \mathbf{R}^2 . However, it will be clear that the same proof works equally well in every dimension.

So from now on $n = 2$ and the goal is to prove the corresponding (3). The operators $C_{2,d,A}$ that we would like to understand, are given by

$$C_{2,d,A}f(x) = p.v. \int_{\mathbf{R}^2} f(x+t) \left(\frac{\Delta_{t_1}^{(1)}}{t_1} \circ \frac{\Delta_{t_2}^{(2)}}{t_2} A(x) \right)^d \frac{dt_1}{t_1} \frac{dt_2}{t_2}. \quad (4)$$

¹This reduction is a simple consequence of the fact that if one writes the analytic function F as a power series, the generic operator $C_{n,F,A}$ itself becomes a series involving all the commutators $C_{n,d,A}$. The polynomial bounds are necessary for this series to be absolutely convergent.

If $a := \frac{\partial^2 A}{\partial x_1 \partial x_2}$, then one observes that

$$\frac{\Delta_{t_1}^{(1)}}{t_1} \circ \frac{\Delta_{t_2}^{(2)}}{t_2} A(x) = \int_{[0,1]^2} a(x_1 + \alpha t_1, x_2 + \beta t_2) d\alpha d\beta. \quad (5)$$

As in [20], using (5) d times, one can see that if a and f are Schwartz functions, the implicit limit in (4) exists and can be rewritten as

$$\int_{\mathbf{R}^{2d+2}} m_{2,d}(\xi, \xi_1, \dots, \xi_d, \eta, \eta_1, \dots, \eta_d) \widehat{f}(\xi, \eta) \widehat{a}(\xi_1, \eta_1) \dots \widehat{a}(\xi_d, \eta_d) \cdot \\ \cdot e^{2\pi i(x_1, x_2) \cdot [(\xi, \eta) + (\xi_1, \eta_1) + \dots + (\xi_d, \eta_d)]} d\xi d\xi_1 \dots d\xi_d d\eta d\eta_1 \dots d\eta_d \quad (6)$$

where

$$m_{2,d}(\xi, \xi_1, \dots, \xi_d, \eta, \eta_1, \dots, \eta_d) := m_{1,d}(\xi, \xi_1, \dots, \xi_d) \cdot m_{1,d}(\eta, \eta_1, \dots, \eta_d) \quad (7)$$

with $m_{1,d}(\xi, \xi_1, \dots, \xi_d)$ and $m_{1,d}(\eta, \eta_1, \dots, \eta_d)$ given by

$$\int_{[0,1]^d} \text{sgn}(\xi + \alpha_1 \xi_1 + \dots + \alpha_d \xi_d) d\alpha_1 \dots d\alpha_d$$

and

$$\int_{[0,1]^d} \text{sgn}(\eta + \beta_1 \eta_1 + \dots + \beta_d \eta_d) d\beta_1 \dots d\beta_d$$

respectively. Because of the formula (6) $C_{2,d}$ can be seen as a $(d+1)$ -linear operator. However, it is important to realize (as in [20]) that even though its symbol $m_{2,d}$ has the nice *product structure* in (7), it is not a classical *bi-parameter* symbol, since $m_{1,d}$ itself is not a classical Marcinkiewicz Hörmander Mihlin multiplier². As a consequence of this fact, the general *polydisc Coifman Meyer theorem* proved in [21], [22] cannot be applied in this case. The strategy would be to combine the techniques of [21], [22] with the new ideas of [19], [20] and to show that (together with some other logarithmical estimates that will be proved in this paper) they are enough to obtain the polynomial bounds of (3). Given these remarks, it would clearly be of great help for the reader, to be already familiar with our earlier arguments in [19], [20].

We will prove the following

Theorem 1.2. *Let $1 < p_1, \dots, p_{d+1} \leq \infty$ and $1 \leq p < \infty$ be so that $1/p_1 + \dots + 1/p_{d+1} = 1/p$. Denote by l the number of indices i for which $p_i \neq \infty$. The operator $C_{2,d}$ extends naturally as a $(d+1)$ -linear operator bounded from $L^{p_1} \times \dots \times L^{p_{d+1}} \rightarrow L^p$ with an operatorial bound of type*

$$C(d) \cdot C(l) \cdot C(p_1) \cdot \dots \cdot C(p_{d+1}) \quad (8)$$

where $C(d)$ grows at most polynomially in d and $C(p_i) = 1$ as long as $p_i = \infty$ for $1 \leq i \leq d+1$.

² $m_{1,d}$ is of course the symbol of the one dimensional d th Calderón commutator [20].

The above Theorem 1.2 is the bi-parameter extension of the corresponding Theorem 1.1 in [20]. If we assume it for a moment, we see that (3) follows from it by taking $p_1 = p$ and $p_2 = \dots = p_{d+1} = \infty$.

To show Theorem 1.2 we will prove that for every $1 \leq i \leq d + 2$ and for every $\phi_1, \dots, \phi_{d+1}$ Schwartz functions, one has

$$\|C_{2,d}^{*i}(\phi_1, \dots, \phi_{d+1})\|_p \leq C(d) \cdot C(l) \cdot C(p_1) \cdot \dots \cdot C(p_{d+1}) \cdot \|\phi_1\|_{p_1} \cdot \dots \cdot \|\phi_{d+1}\|_{p_{d+1}} \quad (9)$$

where $(p_j)_{j=1}^{d+1}$ and p are as before and $(C_{2,d}^{*i})_{i=1}^{d+2}$ are the adjoints of the multilinear operator $C_{2,d}$.³ Standard density and duality arguments as in [20], allow then one to conclude that the estimates in (9) can be naturally extended to arbitrary products of L^{p_j} and L^∞ spaces.⁴

Our plan for the rest of the paper is as follows. In the next section, Section 2, we describe some discrete model operators whose analysis will play an important role in understanding (9). In Section 3 we prove that the main estimates (9) can be reduced to a general theorem for the model operators. In Section 4 we prove the theorem for the discrete model operators of Section 2. In Section 5 we show logarithmical bounds for some shifted Hardy-Littlewood-Paley hybrid operators, which appear naturally in the study of the previous discrete models. Finally, in Section 6 we describe various generalizations of the main Theorem 1.1.

Acknowledgements: The present work has been partially supported by the NSF.

2. DISCRETE MODEL OPERATORS

As mentioned earlier, the main task here would be to describe some discrete model operators, whose analysis is deeply related to the analysis on (9). Because of the formula (6), we now know that

$$C_{2,d} = C_{1,d} \otimes C_{1,d} \quad (10)$$

and so one should not be at all surprised, to find out that these *bi-parameter* model operators that will be introduced, are in fact tensor products of the *one-parameter* discrete model operators of [20]. And also as in [20], these operators are not going to be $(d + 1)$ -linear, but l -linear instead, for some $1 \leq l \leq d + 1$. The explanation for this is similar to the one in [20]. To be able to prove (9), one first decomposes $C_{2,d}$ into polynomially (in d) many *bi-parameter paraproduct like pieces* and then estimate each such piece independently on d . To be able to achieve this, one has first to realize that one can estimate most of the L^∞ functions easily by their L^∞ norms and reduce (9) in this way to the corresponding estimate for some *minimal* l -linear operators. To prove the desired bounds for these minimal operators, one has to interpolate between some Banach and quasi-Banach estimates, as in [20]. The Banach estimates are easy, but the quasi-Banach estimates are hard. One has to discretize the operators carefully, in order to understand them completely. And this is (in a few words) how one arrives at the model operators. Their definition is as follows.

A smooth function $\Phi(x)$ of one variable is said to be a bump function adapted to a dyadic interval I , if and only if one has

³For symmetry, we also use the notation $C_{2,d} = C_{2,d}^{*d+2}$.

⁴The reader is also referred to our earlier [20] for an explanation of why does one need the larger set of estimates in (9) for $C_{2,d}$ and its adjoints, even though one is interested in the more particular (3).

$$|\partial^\alpha \Phi(x)| \lesssim \frac{1}{|I|^\alpha} \frac{1}{\left(1 + \frac{\text{dist}(x, I)}{|I|}\right)^M}$$

for all derivatives α satisfying $|\alpha| \leq 5$ and any large $M > 0$ with the implicit constants depending on it. Then, if $1 \leq q \leq \infty$, we say that $|I|^{-1/q} \Phi$ is an L^q normalized bump adapted to I . The function $\Phi(x)$ is said to be of Ψ type if $\int_{\mathbf{R}} \Phi(x) dx = 0$, otherwise is said to be of Φ type.

A smooth function $\Phi(x, y)$ of two variables is said to be a bump function adapted to the dyadic rectangle $R = I \times J$ if and only if it is of the form $\Phi(x, y) = \Phi_1(x) \cdot \Phi_2(y)$ with $\Phi_1(x)$ adapted to I and $\Phi_2(y)$ adapted to J . If I is a dyadic interval and \mathbf{n} an integer, we denote by $I_{\mathbf{n}} := I + \mathbf{n}|I|$ the dyadic interval having the same length as I but sitting \mathbf{n} units of length $|I|$ away from it.

Fix now $1 \leq l \leq d + 1$ and $\mathbf{n}_1 = (\mathbf{n}_1^1, \mathbf{n}_1^2), \dots, \mathbf{n}_l = (\mathbf{n}_l^1, \mathbf{n}_l^2)$ arbitrary pairs of integers. Define also $\mathbf{n}_{l+1} := (0, 0)$. Consider families $(\Phi_{R_{\mathbf{n}_j}}^j)_R$ for $1 \leq j \leq l + 1$ of L^2 normalized bump functions adapted to dyadic rectangles $R_{\mathbf{n}_j} = I_{\mathbf{n}_j^1} \times J_{\mathbf{n}_j^2}$ where $R = I \times J$ runs inside a given finite collection \mathcal{R} of dyadic rectangles in the plane. Assume also that at least two of the families $(\Phi_{I_{\mathbf{n}_j^1}}^j)_I$ for $1 \leq j \leq l + 1$ are of Ψ type and that the same is true for the families $(\Phi_{J_{\mathbf{n}_j^2}}^j)_J$ for $1 \leq j \leq l + 1$.

The discrete model operator associated to these families of functions is defined by

$$T_{\mathcal{R}}(f_1, \dots, f_l) = \sum_{R \in \mathcal{R}} \frac{1}{|R|^{(l-1)/2}} \langle f_1, \Phi_{R_{\mathbf{n}_1}}^1 \rangle \dots \langle f_l, \Phi_{R_{\mathbf{n}_l}}^l \rangle \Phi_R^{l+1}. \quad (11)$$

The following theorem holds.

Theorem 2.1. *For any such a finite family of arbitrary dyadic rectangles, the l -linear operator $T_{\mathcal{R}}$ maps $L^{p_1} \times \dots \times L^{p_l} \rightarrow L^p$ boundedly, for any $1 < p_1, \dots, p_l < \infty$ with $1/p_1 + \dots + 1/p_l = 1/p$ and $0 < p < \infty$, with a bound of type*

$$O\left(\prod_{j=1}^2 \log^2 < \mathbf{n}_1^j > \dots \log^2 < \mathbf{n}_l^j >\right) \quad (12)$$

where in general, $< m >$ simply denotes $2 + |m|$. And the implicit constants are allowed to depend on l .

This theorem is the bi-parameter generalization of Theorem 3.1 in [20]. As pointed out there, standard arguments based on scale invariance and interpolation, allows one to reduce the above Theorem 2.1 to the more precise statement that for every $f_j \in L^{p_j}$ with $\|f_j\|_{p_j} = 1$ and measurable set $E \subseteq \mathbf{R}^2$ of measure 1, there exists a subset $E' \subseteq E$ with $|E'| \sim |E|$ so that

$$\sum_{R \in \mathcal{R}} \frac{1}{|R|^{(l-1)/2}} |\langle f_1, \Phi_{R_{\mathbf{n}_1}}^1 \rangle| \dots |\langle f_l, \Phi_{R_{\mathbf{n}_l}}^l \rangle| |\langle f_{l+1}, \Phi_R^{l+1} \rangle| \lesssim \prod_{j=1}^2 \log^2 < \mathbf{n}_1^j > \dots \log^2 < \mathbf{n}_l^j > \quad (13)$$

where $f_{l+1} := \chi_{E'}$. As in [20], the fact that one loses only logarithmical bounds in the above estimates, will be of a crucial importance later on.

3. REDUCTION TO THE MODEL OPERATORS

The goal of this section is to show that indeed (9) can be reduced to Theorem 2.1 or more precisely to its weaker but more precise variant (13). In particular, one can find here a description of all the ideas that are necessary to understand why it is possible to estimate the biparameter Calderón commutators $C_{2,d}$ with bounds that grow at most polynomially in d .

The reader familiar with our previous work will realize that this section is in fact a *tensor product* of the corresponding section in [20] with itself. As there, the first task is to decompose $C_{2,d}$ into polynomially many *biparameter paraproduct like pieces* which will be studied later on.

Non-compact and compact Littlewood-Paley decompositions. Let $\Phi(x)$ be a Schwartz function which is even, positive and satisfying $\int_{\mathbf{R}} \Phi(x) dx = 1$. Define also $\Psi(x)$ by

$$\Psi(x) = \Phi(x) - \frac{1}{2}\Phi\left(\frac{x}{2}\right)$$

and observe that $\int_{\mathbf{R}} \Psi(x) dx = 0$.

Then, as always, consider the functions $\Psi_k(x)$ and $\Phi_k(x)$ defined by $2^k\Psi(2^kx)$ and $2^k\Phi(2^kx)$ respectively, for every integer $k \in \mathbf{Z}$. Notice also that all the L^1 norms of Φ_k are equal to 1. Since $\Psi_k(x) = \Phi_k(x) - \Phi_{k-1}(x)$ one can see that

$$\sum_{k \leq k_0} \Psi_k = \Phi_{k_0}$$

and so

$$\sum_{k \in \mathbf{Z}} \Psi_k = \delta_0$$

or equivalently

$$\sum_{k \in \mathbf{Z}} \widehat{\Psi}_k(\xi) = 1 \tag{14}$$

for almost every $\xi \in \mathbf{R}$. On the other hand, as observed in [20], since $\widehat{\Psi}(0) = \widehat{\Psi}'(0) = 0$ one can write $\widehat{\Psi}(\xi)$ as

$$\widehat{\Psi}(\xi) = \xi^2 \varphi(\xi)$$

for some other smooth and rapidly decaying function φ .

These are what we called the non-compact (in frequency) Littlewood-Paley decompositions. The compact ones are obtained similarly, the only difference being that instead of considering the Schwartz function Φ before, one starts with another one having the property that $\text{supp } \widehat{\Phi} \subseteq [-1, 1]$ and $\widehat{\Phi}(0) = 1$.

As explained in [20], the advantage of the non-compact Littlewood-Paley projections is reflected in the perfect estimate

$$|f * \Phi_k(x)| \leq \|f\|_{\infty} \tag{15}$$

which plays an important role in the argument.

The generic decomposition of $C_{2,d}$. Using (6), if f, f_1, \dots, f_{d+1} are all Schwartz functions, one can write the $(d+2)$ -linear form associated to $C_{2,d}$ as

$$\begin{aligned} & \int_{\substack{\xi+\xi_1+\dots+\xi_{d+1}=0 \\ \eta+\eta_1+\dots+\eta_{d+1}=0}} \left(\int_{[0,1]^d} \text{sgn}(\xi + \alpha_1\xi_1 + \dots + \alpha_d\xi_d) d\alpha_1 \dots d\alpha_d \right) \cdot \\ & \left(\int_{[0,1]^d} \text{sgn}(\eta + \beta_1\eta_1 + \dots + \beta_d\eta_d) d\beta_1 \dots d\beta_d \right) \cdot \\ & \widehat{f}(\xi, \eta) \widehat{f_1}(\xi_1, \eta_1) \dots \widehat{f_{d+1}}(\xi_{d+1}, \eta_{d+1}) d\xi d\xi_1 \dots d\xi_{d+1} d\eta d\eta_1 \dots d\eta_{d+1}. \end{aligned} \quad (16)$$

Then, by using the Littlewood-Paley decompositions in (14) several times one can write

$$1 = \sum_{l_0, l_1, \dots, l_{d+1} \in \mathbb{Z}} \widehat{\Psi}_{l_0}(\xi) \widehat{\Psi}_{l_1}(\xi_1) \dots \widehat{\Psi}_{l_d}(\xi_d) \widehat{\Psi}_{l_{d+1}}(\xi_{d+1}). \quad (17)$$

As in [20] since for every $(d+2)$ tuple $(l_0, l_1, \dots, l_{d+1}) \in \mathbb{Z}^{d+2}$ one has that either $l_0 \geq l_1, \dots, l_{d+1}$ or $l_1 \geq l_0, \dots, l_{d+1}$... or $l_{d+1} \geq l_0, \dots, l_d$, fixing always the biggest parameter and summing over the rest of them, one can rewrite (17) as

$$\begin{aligned} & \sum_l \widehat{\Psi}_l(\xi) \widehat{\Phi}_l(\xi_1) \dots \widehat{\Phi}_l(\xi_d) \widehat{\Phi}_l(\xi_{d+1}) + \\ & \dots + \\ & \sum_l \widehat{\Phi}_l(\xi) \widehat{\Phi}_l(\xi_1) \dots \widehat{\Phi}_l(\xi_d) \widehat{\Psi}_l(\xi_{d+1}). \end{aligned} \quad (18)$$

Also as in [20], we use in (18) compact Littlewood-Paley decompositions for the ξ and ξ_{d+1} variables and non-compact ones for the rest of them. Every single term in the decomposition (18) contains only one Ψ type of a function and we would like to have (at least) two. To be able to producte another one, one has to recall that $\xi + \xi_1 + \dots + \xi_{d+1} = 0$. Taking this into account, let us take a look at the second (for instance) term in (18) in the particular case when $l = 0$. We rewrite it for simplicity as

$$\widehat{\Phi}(\xi) \widehat{\Psi}(\xi_1) \dots \widehat{\Phi}(\xi_d) \widehat{\Phi}(\xi_{d+1}). \quad (19)$$

We know from before that $\widehat{\Psi}(\xi_1) = \xi_1^2 \widehat{\varphi}(\xi_1)$ and so we can write

$$\widehat{\Psi}(\xi_1) = \xi_1 \widehat{\varphi}(\xi_1) (-\xi - \xi_2 - \dots - \xi_{d+1}) = -\xi_1 \xi \widehat{\varphi}(\xi_1) - \xi_1 \xi_2 \widehat{\varphi}(\xi_1) - \dots - \xi_1 \xi_{d+1} \widehat{\varphi}(\xi_1).$$

Using this in (19) allows one to decompose it as another sum of $O(d)$ terms, containing this time two function of Ψ type, since besides $\xi_1 \widehat{\varphi}(\xi_1)$ one finds now either a factor of type $\xi \widehat{\Phi}(\xi)$ or of type $\xi_j \widehat{\Phi}_j(\xi_j)$ for some $j = 2, \dots, d+1$.

If one performs a similar decomposition for every scale $l \in \mathbb{Z}$ and each of the terms in (18) one obtains a splitting of the function $1_{\{\xi+\xi_1+\dots+\xi_{d+1}=0\}}$ as a sum of $O(d^2)$ expressions whose generic inner terms contain two functions of Ψ type as desired.

Since we are this time in the biparameter setting, one has to decompose $1_{\{\eta+\eta_1+\dots+\eta_{d+1}=0\}}$ in a completely similar manner. Combining these two decompositions, allows us to rewrite the $(d+2)$ linear form of $C_{2,d}$ as

$$\begin{aligned}
& \sum_{k_1, k_2 \in \mathbb{Z}} \int_{\substack{\xi + \xi_1 + \dots + \xi_{d+1} = 0 \\ \eta + \eta_1 + \dots + \eta_{d+1} = 0}} \left(\int_{[0,1]^d} \text{sgn}(\xi + \alpha_1 \xi_1 + \dots + \alpha_d \xi_d) d\alpha_1 \dots d\alpha_d \right) \cdot \\
& \left(\int_{[0,1]^d} \text{sgn}(\eta + \beta_1 \eta_1 + \dots + \beta_d \eta_d) d\beta_1 \dots d\beta_d \right) \cdot \\
& \widehat{\Phi}_{k_1}^{1,0}(\xi) \widehat{\Phi}_{k_1}^{1,1}(\xi_1) \dots \widehat{\Phi}_{k_1}^{1,d}(\xi_d) \widehat{\Phi}_{k_1}^{1,d+1}(\xi_{d+1}) \cdot \\
& \widehat{\Phi}_{k_2}^{2,0}(\eta) \widehat{\Phi}_{k_2}^{2,1}(\eta_1) \dots \widehat{\Phi}_{k_2}^{2,d}(\eta_d) \widehat{\Phi}_{k_2}^{2,d+1}(\eta_{d+1}) \cdot \\
& \widehat{f}(\xi, \eta) \widehat{f}_1(\xi_1, \eta_1) \dots \widehat{f}_{d+1}(\xi_{d+1}, \eta_{d+1}) d\xi d\xi_1 \dots d\xi_{d+1} d\eta d\eta_1 \dots d\eta_{d+1},
\end{aligned} \tag{20}$$

which completes our generic decomposition.

Recall that at least two of the families $(\widehat{\Phi}_{k_1}^{1,j}(\xi_j))_{k_1}$ for $0 \leq j \leq d+1$ are of Ψ type and likewise at least two of the families $(\widehat{\Phi}_{k_2}^{2,j}(\eta_j))_{k_2}$ for $0 \leq j \leq d+1$ are of Ψ type as well. We denote those indices by i_1, i_2 and j_1, j_2 respectively. There are several cases that one has to consider which correspond to the positions of these indices. We call an index *intermediate* if it is between 1 and d and *extremal* if it is either 0 or $d+1$. In [19] we essentially witnessed two cases. Case 1 was when at least one of the Ψ positions corresponded to an intermediate index and Case 2 was when both of the Ψ positions were extremal. Since we now work in the biparameter setting, there are as a consequence four possible cases of type Case $i \otimes$ Case j for $1 \leq i, j \leq 2$.

Case 1 \otimes Case 1. Assume here that $i_1 = j_1 = 0$ and $i_2 = j_2 = 1$. As mentioned earlier, the fact that $i_1 = i_2 = 0$ is not important, they can be anywhere else in the interval $[0, d+1]$. Also, the fact that the intermediate indices i_2 and j_2 have been chosen to be equal is not important either, but we chose them so for the simplicity of the notation. As in [20] we would like now to expand the two implicit symbols in (20).

As there, let us denote by $\widetilde{\xi} := \xi + \alpha_2 \xi_2 + \dots + \alpha_d \xi_d$ and by $\widetilde{\eta} := \eta + \beta_2 \eta_2 + \dots + \beta_d \eta_d$ and recall from [20] that the idea is to treat the first symbol of (20) as being dependent on the variables ξ_1 and $\widetilde{\xi}$ and similarly the second symbol of (20) as being dependent on η_1 and $\widetilde{\eta}$. Also, since most of our functions do not have compact support in frequency, we need to consider some other compact Littlewood-Paley decompositions. We first write as in [20]

$$1 = \sum_{l_0, l_1} \widehat{\Psi}_{l_0}(\widetilde{\xi}) \widehat{\Psi}_{l_1}(\xi_1) = \sum_{l_0 < l_1} \dots + \sum_{l_0 \sim l_1} \dots + \sum_{l_0 > l_1} \dots$$

which can be rewritten as

$$\sum_{r_1} \widehat{\Phi}_{r_1}(\widetilde{\xi}) \widehat{\Psi}_{r_1}(\xi_1) + \sum_{r_1} \widehat{\Psi}_{r_1}(\widetilde{\xi}) \widehat{\Phi}_{r_1}(\xi_1) + \sum_{r_1} \widehat{\Psi}_{r_1}(\widetilde{\xi}) \widehat{\Phi}_{r_1}(\xi_1). \tag{21}$$

Then, we consider an identical decomposition, but for the variables $\widetilde{\eta}$ and η_1 this time, where the summation is indexed over the parameter r_2 . If we insert (21) into (20) it becomes a sum of three distinct expressions that generate the subcases 1_a , 1_b and 1_c respectively. If in addition one inserts the analogous formula of (21) for the variables $\widetilde{\eta}$ and η_1 into (20) as well, one ends up with nine *biparameter subcases* of type Case $1_a \otimes$ Case 1_a , Case $1_a \otimes$ Case 1_b and so on.

Case $1_a \otimes \text{Case } 1_a$. To analyze the impact that these extra decompositions have, we consider for simplicity (as in [20]) the particular term corresponding to $k_1 = k_2 = 0$. However, the argument we use is scale invariant.

Let us ignore the symbol in (20) for now and just concentrate on the remaining expression which becomes

$$\begin{aligned}
 & \left(\sum_{r_1} [\widehat{\Phi}_{r_1}(\tilde{\xi}) \widehat{\Psi}_{r_1}(\xi_1)] \widehat{\Phi}_0^{1,0}(\xi) \widehat{\Phi}_0^{1,1}(\xi_1) \dots \widehat{\Phi}_0^{1,d}(\xi_d) \widehat{\Phi}_0^{1,d+1}(\xi_{d+1}) \right) \cdot \\
 & \left(\sum_{r_2} [\widehat{\Phi}_{r_2}(\tilde{\eta}) \widehat{\Psi}_{r_2}(\eta_1)] \widehat{\Phi}_0^{2,0}(\eta) \widehat{\Phi}_0^{2,1}(\eta_1) \dots \widehat{\Phi}_0^{2,d}(\eta_d) \widehat{\Phi}_0^{2,d+1}(\eta_{d+1}) \right) = \\
 & \left(\sum_{r_1 \leq 0} \dots + \sum_{r_1 > 0} \dots \right) \cdot \\
 & \left(\sum_{r_2 \leq 0} \dots + \sum_{r_2 > 0} \dots \right) := \\
 & (1'_a + 1''_a) \otimes (1'_a + 1''_a)
 \end{aligned} \tag{22}$$

which allows us to split our existing subcase into four additional subcases.

Case $1'_a \otimes \text{Case } 1'_a$. So this corresponds to the situation when both r_1 and r_2 are negative. As in [20], using the fact that $\widehat{\Psi}_{r_1}(\xi_1)$ is compactly supported and given that $\widehat{\Phi}_0^{1,1}(\xi_1)$ is also of Ψ type (in fact it is of the form $\xi_1 \widehat{\varphi}(\xi_1)$) one can rewrite *the ξ part* of (22) as

$$\begin{aligned}
 & \sum_{r_1 \leq 0} 2^{r_1} \widehat{\Phi}_{r_1}(\tilde{\xi}) \widehat{\Phi}_0^{1,0}(\xi) \widehat{\Psi}_{r_1}^{1,1}(\xi_1) \dots \widehat{\Phi}_0^{1,d}(\xi_d) \widehat{\Phi}_0^{1,d+1}(\xi_{d+1}) = \\
 & \sum_{r_1 \leq 0} 2^{r_1} \left[\widehat{\Phi}_{r_1}(\tilde{\xi}) \widehat{\Psi}_{r_1}^{1,1}(\xi_1) \right] \cdot \widehat{\Phi}_0^{1,0}(\xi) \widehat{\Psi}_{r_1}^{1,1}(\xi_1) \dots \widehat{\Phi}_0^{1,d+1}(\xi_{d+1}) \widehat{\Phi}_{r_1}(\tilde{\xi})
 \end{aligned}$$

for naturally chosen compactly supported functions $\widehat{\Phi}_{r_1}(\tilde{\xi})$, $\widehat{\Psi}_{r_1}^{1,1}(\xi_1)$ and $\widehat{\Psi}_{r_1}^{1,1}(\xi_1)$.

This allows us to split the symbol

$$\left(\int_0^1 \text{sgn}(\tilde{\xi} + \alpha_1 \xi_1) d\alpha_1 \right) \widehat{\Phi}_{r_1}(\tilde{\xi}) \widehat{\Psi}_{r_1}^{1,1}(\xi_1)$$

as a double Fourier series of the form

$$\sum_{\tilde{n}, \tilde{n}_1 \in \mathbb{Z}} C_{\tilde{n}, \tilde{n}_1}^{r_1} e^{2\pi i \frac{\tilde{n}}{2^{r_1}} \tilde{\xi}} \cdot e^{2\pi i \frac{\tilde{n}_1}{2^{r_1}} \xi_1} \tag{23}$$

where the Fourier coefficients satisfy the quadratic estimates

$$\left| C_{\tilde{n}, \tilde{n}_1}^{r_1} \right| = \left| C_{\tilde{n}, \tilde{n}_1} \right| \lesssim \frac{1}{<\tilde{n}>^2} \frac{1}{<\tilde{n}_1>^\#} \tag{24}$$

for an arbitrarily large number $\# > 0$. See [20] for these important estimates.

Clearly, there are similar calculations that one can make for *the η part* of (22). Using both of them, one can see that the particular contribution of $1'_a \otimes 1'_a$ in (20) (at scale 1) becomes

$$\begin{aligned}
& \int_{[0,1]^{d-1}} \int_{[0,1]^{d-1}} \sum_{r_1 \leq 0} 2^{r_1} \sum_{r_2 \leq 0} 2^{r_2} \sum_{\widetilde{n}, \widetilde{n}_1} C_{\widetilde{n}, \widetilde{n}_1}^{r_1} \sum_{\widetilde{\widetilde{n}}, \widetilde{\widetilde{n}}_1} C_{\widetilde{\widetilde{n}}, \widetilde{\widetilde{n}}_1}^{r_2} \int_{\substack{\xi + \xi_1 + \dots + \xi_{d+1} = 0 \\ \eta + \eta_1 + \dots + \eta_{d+1} = 0}} \\
& \quad \left[\widehat{\Phi}_0^{1,0}(\xi) e^{2\pi i \frac{\widetilde{n}}{2^{r_1}} \xi} \cdot \widehat{\Phi}_0^{2,0}(\eta) e^{2\pi i \frac{\widetilde{\widetilde{n}}}{2^{r_2}} \eta} \right] \cdot \\
& \quad \left[\widehat{\Psi}_{r_1}^{1,1}(\xi_1) e^{2\pi i \frac{\widetilde{n}_1}{2^{r_1}} \xi_1} \cdot \widehat{\Psi}_{r_2}^{2,1}(\eta_1) e^{2\pi i \frac{\widetilde{\widetilde{n}}_1}{2^{r_2}} \eta_1} \right] \cdot \\
& \quad \left[\widehat{\Phi}_0^{1,2}(\xi_2) e^{2\pi i \frac{\widetilde{n}}{2^{r_1}} \alpha_2 \xi_2} \cdot \widehat{\Phi}_0^{2,2}(\eta_2) e^{2\pi i \frac{\widetilde{\widetilde{n}}}{2^{r_2}} \beta_2 \eta_2} \right] \cdot \\
& \quad \dots \\
& \quad \left[\widehat{\Phi}_0^{1,d}(\xi_d) e^{2\pi i \frac{\widetilde{n}}{2^{r_1}} \alpha_d \xi_d} \cdot \widehat{\Phi}_0^{2,d}(\eta_d) e^{2\pi i \frac{\widetilde{\widetilde{n}}}{2^{r_2}} \beta_d \eta_d} \right] \cdot \\
& \quad \left[\widehat{\Phi}_0^{1,d+1}(\xi_{d+1}) \cdot \widehat{\Phi}_0^{2,d+1}(\eta_{d+1}) \right] \cdot \\
& \quad \left[\widehat{\Phi}_{r_1}(\widetilde{\xi}) \cdot \widehat{\Phi}_{r_2}(\widetilde{\eta}) \right] \cdot
\end{aligned} \tag{25}$$

$$\widehat{f}(\xi, \eta) \widehat{f}_1(\xi_1, \eta_1) \dots \widehat{f}_{d+1}(\xi_{d+1}, \eta_{d+1}) d\xi d\xi_1 \dots d\xi_{d+1} d\eta d\eta_1 \dots d\eta_{d+1} d\alpha_2 \dots d\alpha_d d\beta_2 \dots d\beta_d.$$

Now, if one fixes $\vec{\alpha}, \vec{\beta}, r_1, r_2, \widetilde{n}, \widetilde{n}_1, \widetilde{\widetilde{n}}, \widetilde{\widetilde{n}}_1$, the corresponding inner expression in (25) becomes

$$\begin{aligned}
& \int_{\substack{\xi + \xi_1 + \dots + \xi_{d+1} = 0 \\ \eta + \eta_1 + \dots + \eta_{d+1} = 0}} \\
& \quad \left[\widehat{f}(\xi, \eta) \cdot \widehat{\Phi}_0^{1,0}(\xi) e^{2\pi i \frac{\widetilde{n}}{2^{r_1}} \xi} \cdot \widehat{\Phi}_0^{2,0}(\eta) e^{2\pi i \frac{\widetilde{\widetilde{n}}}{2^{r_2}} \eta} \right] \cdot \\
& \quad \left[\widehat{f}_1(\xi_1, \eta_1) \cdot \widehat{\Psi}_{r_1}^{1,1}(\xi_1) e^{2\pi i \frac{\widetilde{n}_1}{2^{r_1}} \xi_1} \cdot \widehat{\Psi}_{r_2}^{2,1}(\eta_1) e^{2\pi i \frac{\widetilde{\widetilde{n}}_1}{2^{r_2}} \eta_1} \right] \cdot \\
& \quad \dots \\
& \quad \left[\widehat{f}_d(\xi_d, \eta_d) \cdot \widehat{\Phi}_0^{1,d}(\xi_d) e^{2\pi i \frac{\widetilde{n}}{2^{r_1}} \alpha_d \xi_d} \cdot \widehat{\Phi}_0^{2,d}(\eta_d) e^{2\pi i \frac{\widetilde{\widetilde{n}}}{2^{r_2}} \beta_d \eta_d} \right] \cdot \\
& \quad \left[\widehat{f}_{d+1}(\xi_{d+1}, \eta_{d+1}) \cdot \widehat{\Phi}_0^{1,d+1}(\xi_{d+1}) \cdot \widehat{\Phi}_0^{2,d+1}(\eta_{d+1}) \right] \cdot \\
& \quad \widehat{\Phi}_{r_1}(\xi + \alpha_2 \xi_2 + \dots + \alpha_d \xi_d) \cdot \widehat{\Phi}_{r_2}(\eta + \beta_2 \eta_2 + \dots + \beta_d \eta_d) \\
& \quad d\xi d\xi_1 \dots d\xi_{d+1} d\eta d\eta_1 \dots d\eta_{d+1}.
\end{aligned} \tag{26}$$

To be able to continue the calculations we need the following lemma.

Lemma 3.1. *If $F, F_1, \dots, F_{d+1}, \widetilde{\Phi}, \widetilde{\widetilde{\Phi}}$ are all Schwartz functions, then one has*

$$\int_{\substack{\xi+\xi_1+\dots+\xi_{d+1}=0 \\ \eta+\eta_1+\dots+\eta_{d+1}=0}} \widehat{F}(\xi, \eta) \widehat{F}_1(\xi_1, \eta_1) \dots \widehat{F}_{d+1}(\xi_{d+1}, \eta_{d+1}) \cdot \\ \widetilde{\widetilde{\Phi}}(a\xi + a_1\xi_1 + \dots + a_{d+1}\xi_{d+1}) \cdot \widetilde{\widetilde{\Phi}}(b\eta + b_1\eta_1 + \dots + b_{d+1}\eta_{d+1}) d\xi d\xi_1 \dots d\xi_{d+1} d\eta d\eta_1 \dots d\eta_{d+1} = \\ \int_{\mathbf{R}^4} F(x_1 - at_1, x_2 - at_2) F_1(x_1 - a_1t_1, x_2 - b_1t_2) \dots F_{d+1}(x_1 - a_{d+1}t_1, x_2 - b_{d+1}t_2) \widetilde{\Phi}(t_1) \widetilde{\widetilde{\Phi}}(t_2) dx_1 dx_2 dt_1 dt_2$$

for $a, a_1, \dots, a_{d+1}, b, b_1, \dots, b_{d+1}$ arbitrary real numbers.

This Lemma 3.1 is the biparameter extension of Lemma 4.1 in [20] and since its proof doesn't use any new ideas it is left to the reader. As pointed out in [20] there is also a natural generalization of it, which states that the formula works for more than two averages (so one can take an arbitrary number of $\widetilde{\Phi}$ functions and another arbitrary number of $\widetilde{\widetilde{\Phi}}$ ones).

As in [20], if G is now an arbitrary Schwartz function and a a real number we denote by G^a the function defined by

$$\widehat{G^a}(\xi) = \widehat{G}(\xi) e^{2\pi i a \xi}.$$

Alternatively, one has $G^a(x) = G(x - a)$. Using the above lemma and this notation the previous (26) can be rewritten as

$$\int_{\mathbf{R}^4} \left(f * \Phi_0^{1,0,\frac{\bar{n}}{2^r1}} \otimes \Phi_0^{2,0,\frac{\bar{n}}{2^r2}} \right) (x_1 - t_1, x_2 - t_2) \cdot \\ \left(f_1 * \Psi_{r_1}^{1,1,\frac{\bar{n}_1}{2^r1}} \otimes \Psi_{r_2}^{2,1,\frac{\bar{n}_1}{2^r2}} \right) (x_1, x_2) \cdot \\ \prod_{j=2}^d \left(f_j * \Phi_0^{1,j,\frac{\bar{n}}{2^r1}\alpha_j} \otimes \Phi_0^{2,j,\frac{\bar{n}}{2^r2}\beta_j} \right) (x_1 - \alpha_j t_1, x_2 - \beta_j t_2) \cdot \\ \left(f_{d+1} * \Phi_0^{1,d+1} \otimes \Phi_0^{2,d+1} \right) (x_1, x_2) \cdot \\ \Phi_{r_1}(t_1) \Phi_{r_2}(t_2) dt_1 dt_2 dx_1 dx_2 = \\ \int_{\mathbf{R}^4} \left(f * \Phi_0^{1,0,\frac{\bar{n}}{2^r1}} \otimes \Phi_0^{2,0,\frac{\bar{n}}{2^r2}} \right) \left(x_1 - \frac{t_1}{2^{r_1}}, x_2 - \frac{t_2}{2^{r_2}} \right) \cdot \\ \left(f_1 * \Psi_{r_1}^{1,1,\frac{\bar{n}_1}{2^r1}} \otimes \Psi_{r_2}^{2,1,\frac{\bar{n}_1}{2^r2}} \right) (x_1, x_2) \cdot \\ \prod_{j=2}^d \left(f_j * \Phi_0^{1,j,\frac{\bar{n}}{2^r1}\alpha_j} \otimes \Phi_0^{2,j,\frac{\bar{n}}{2^r2}\beta_j} \right) \left(x_1 - \frac{\alpha_j t_1}{2^{r_1}}, x_2 - \frac{\beta_j t_2}{2^{r_2}} \right) \cdot \\ \left(f_{d+1} * \Phi_0^{1,d+1} \otimes \Phi_0^{2,d+1} \right) (x_1, x_2) \cdot$$

$$\begin{aligned}
& \Phi_0(t_1)\Phi_0(t_2)dt_1dt_2dx_1dx_2 = \\
& \int_{\mathbf{R}^4} \left(f * \Phi_0^{1,0,\frac{\bar{n}-t_1}{2^{r_1}}} \otimes \Phi_0^{2,0,\frac{\bar{n}-t_2}{2^{r_2}}} \right) (x_1, x_2) \cdot \\
& \left(f_1 * \Psi_{r_1}^{1,1,\frac{\bar{n}_1}{2^{r_1}}} \otimes \Psi_{r_2}^{2,1,\frac{\bar{n}_1}{2^{r_2}}} \right) (x_1, x_2) \cdot \\
& \prod_{j=2}^d \left(f_j * \Phi_0^{1,j,\frac{\bar{n}-t_1}{2^{r_1}}\alpha_j} \otimes \Phi_0^{2,j,\frac{\bar{n}-t_2}{2^{r_2}}\beta_j} \right) (x_1, x_2) \cdot \\
& (f_{d+1} * \Phi_0^{1,d+1} \otimes \Phi_0^{2,d+1}) (x_1, x_2) \cdot \\
& \Phi_0(t_1)\Phi_0(t_2)dt_1dt_2dx_1dx_2.
\end{aligned} \tag{27}$$

We have to remember now that all the calculations so far have been made under the assumption that $k_1 = k_2 = 0$, but they can clearly be performed in general and then the analogous formula of (27) is

$$\begin{aligned}
& \int_{\mathbf{R}^4} \left(f * \Phi_{k_1}^{1,0,\frac{\bar{n}-t_1}{2^{k_1+r_1}}} \otimes \Phi_{k_2}^{2,0,\frac{\bar{n}-t_2}{2^{k_2+r_2}}} \right) (x_1, x_2) \cdot \\
& \left(f_1 * \Psi_{k_1+r_1}^{1,1,\frac{\bar{n}_1}{2^{k_1+r_1}}} \otimes \Psi_{k_2+r_2}^{2,1,\frac{\bar{n}_1}{2^{k_2+r_2}}} \right) (x_1, x_2) \cdot \\
& \prod_{j=2}^d \left(f_j * \Phi_{k_1}^{1,j,\frac{\bar{n}-t_1}{2^{k_1+r_1}}\alpha_j} \otimes \Phi_{k_2}^{2,j,\frac{\bar{n}-t_2}{2^{k_2+r_2}}\beta_j} \right) (x_1, x_2) \cdot \\
& (f_{d+1} * \Phi_{k_1}^{1,d+1} \otimes \Phi_{k_2}^{2,d+1}) (x_1, x_2) \cdot \\
& \Phi_0(t_1)\Phi_0(t_2)dt_1dt_2dx_1dx_2.
\end{aligned} \tag{28}$$

In conclusion, if one denotes by $\vec{\alpha} = (\alpha_2, \dots, \alpha_d)$ and by $\vec{\beta} = (\beta_2, \dots, \beta_d)$ one sees that the part of (20) that corresponds to Case $1'_a \otimes$ Case $1'_a$ can be written as

$$\begin{aligned}
& \int_{[0,1]^{d-1}} \int_{[0,1]^{d-1}} \left(\sum_{r_1 \leq 0} 2^{r_1} \sum_{r_2 \leq 0} 2^{r_2} \sum_{\bar{n}, \bar{n}_1} C_{\bar{n}, \bar{n}_1}^{r_1} \sum_{\bar{n}, \bar{n}_1} C_{\bar{n}, \bar{n}_1}^{r_2} \cdot C_{2,d}^{r_1, r_2, \bar{n}, \bar{n}_1, \bar{n}, \bar{n}_1, \vec{\alpha}, \vec{\beta}, t_1, t_2} \right) \cdot \\
& \Phi_0(t_1)\Phi_0(t_2)dt_1dt_2d\vec{\alpha}d\vec{\beta}
\end{aligned} \tag{29}$$

where $C_{2,d}^{r_1, r_2, \bar{n}, \bar{n}_1, \bar{n}, \bar{n}_1, \vec{\alpha}, \vec{\beta}, t_1, t_2}$ is the operator whose $(d+2)$ linear form is given by summing over k_1, k_2 the inner expressions of (28).

To prove (9) for (29) one would need to prove it for the operators $C_{2,d}^{r_1, r_2, \bar{n}, \bar{n}_1, \bar{n}, \bar{n}_1, \vec{\alpha}, \vec{\beta}, t_1, t_2}$ with bounds that are summable over $r_1, r_2, \bar{n}, \bar{n}_1, \bar{n}, \bar{n}_1$ and integrable over $\vec{\alpha}, \vec{\beta}, t_1, t_2$. It is clear that these operators are essentially biparameter paraproducts and therefore one expects that the

method of [21], [22] should be used. That indeed will be the case, but on the other hand the appearance of all these parameters mentioned earlier, have the role to *shift* the implicit bump functions which appear in their definitions and as a consequence this time one has to be very precise, when evaluates the size of their boundedness constants.

Since in our case away from the indices 0 and 1 all the bump functions are of Φ type, the idea is to use the perfect estimate (15) (or rather, its biparameter variant) to bound all the functions which are in L^∞ . To be more specific, let us denote as in [20] by S the set of all the indices $2 \leq j \leq d$ for which $p_j \neq \infty$. Set $l := |S| + 2$ and freeze all the L^∞ normalized Schwartz functions f_j corresponding to the indices in $\{2, \dots, d\} \setminus S$. The new resulted operator is a *minimal* l -linear operator which will be denoted by $C_{2,d}^{l,r_1,r_2,\tilde{n},\tilde{n}_1,\tilde{n},\tilde{n}_1,\vec{\alpha},\vec{\beta},t_1,t_2}$.

Shifted hybrid maximal and square functions. It is now time to recall a few basic facts about biparameter paraproducts, to be able to go further. Consider two generic families of L^1 normalized bump functions $(\Phi_{k_1}^j)_{k_1}$ and $(\Phi_{k_2}^j)_{k_2}$ for $1 \leq j \leq l+1$ so that in each of them for two indices the corresponding sequences are of Ψ type.

An l -linear biparameter paraproduct is an l -linear operator whose $(l+1)$ -linear form is given by

$$\int_{\mathbf{R}^2} \sum_{k_1, k_2 \in \mathbf{Z}} \prod_{j=1}^{l+1} (f_j * \Phi_{k_1}^j \otimes \Phi_{k_2}^j)(x_1, x_2) dx_1 dx_2. \quad (30)$$

Let us first assume that we are in a case similar to the one considered before and that the Ψ functions appear for the indices $j = 1, 2$. Then, one can estimate the absolute value of (30) by

$$\int_{\mathbf{R}^2} SS(f_1)(x_1, x_2) \cdot SS(f_2)(x_1, x_2) \cdot \prod_{j=2}^{l+1} MM(f_j)(x_1, x_2) dx_1 dx_2$$

where in general $MM(f)(x_1, x_2)$ and $SS(f)(x_1, x_2)$ are defined by

$$MM(f)(x_1, x_2) = \sup_{k_1, k_2} |f * \Phi_{k_1} \otimes \Phi_{k_2}(x_1, x_2)|$$

and

$$SS(f)(x_1, x_2) = \left(\sum_{k_1, k_2} |f * \Phi_{k_1} \otimes \Phi_{k_2}(x_1, x_2)|^2 \right)^{1/2} \quad (31)$$

respectively. In order for (31) to make sense, we assume of course that both $(\Phi_{k_1})_{k_1}$ and $(\Phi_{k_2})_{k_2}$ there, are of Ψ type. Since both MM and SS are known to be bounded in every L^p space for $1 < p < \infty$, the above argument proves that our particular biparameter paraproduct in (30) is bounded from $L^{p_1} \times \dots \times L^{p_l}$ into L^p as long as $1/p_1 + \dots + 1/p_l = 1/p$ and $1 < p_1, \dots, p_l, p < \infty$.

As one can imagine, the above $l^2 \times l^2 \times l^\infty$ argument can be *twisted*, in which case one naturally obtains *hybrid maximal and square functions* of type MS and SM defined by

$$MS(f)(x_1, x_2) = \sup_{k_1} \left(\sum_{k_2} |f * \Phi_{k_1} \otimes \Phi_{k_2}(x_1, x_2)|^2 \right)^{1/2} \quad (32)$$

and

$$SM(f)(x_1, x_2) = \left(\sum_{k_1} \sup_{k_2} |f * \Phi_{k_1} \otimes \Phi_{k_2}(x_1, x_2)|^2 \right)^{1/2} \quad (33)$$

respectively. One has to assume in (32) that the family $(\Phi_{k_2})_{k_2}$ is of Ψ type and that $(\Phi_{k_1})_{k_1}$ is of Ψ type in (33), for both expressions to make sense.

As observed in [22] all these hybrid operators are bounded in L^p for $1 < p < \infty$ as well and as a consequence, one can bound every biparameter paraproduct in arbitrary products of L^p spaces, as long as all of their indices are strictly between 1 and ∞ .

This discussion shows that in order to understand the operator $C_{2,d}^{l,r_1,r_2,\tilde{n},\tilde{n}_1,\tilde{n},\tilde{n}_1,\tilde{\alpha},\tilde{\beta},t_1,t_2}$ (and of course, all the other possible ones) one has to understand how to bound not only the above operators, but also their *shifted* analogs of type $M_{n_1}M_{n_1}$, $S_{n_1}S_{n_2}$, $M_{n_1}S_{n_2}$ and $S_{n_1}M_{n_2}$ which are defined similarly, but with respect to the *shifted* functions $(\Phi_{\frac{n_1}{2^{k_1}}})_{k_1}$ and $(\Phi_{\frac{n_2}{2^{k_2}}})_{k_2}$.

In [19] we understood completely the one-parameter shifted maximal and square functions M_n and S_n and proved their boundedness on L^p spaces with operatorial bounds of type $O(\log < n >)^5$.

Now the arguments of [19] and [22] show that their hybrid biparameter analogs mentioned before, will also be bounded on L^p spaces with operatorial bounds of type $O(\log^2 < n_1 > \log^2 < n_2 >)$, as long as one can prove logarithmic bounds for the so called Fefferman-Stein inequality, namely

$$\left\| \left(\sum_{j=1}^N |M_n f_j|^2 \right)^{1/2} \right\|_p \leq C_p \log^2 < n > \cdot \left\| \left(\sum_{j=1}^N |f_j|^2 \right)^{1/2} \right\|_p \quad (34)$$

which should hold true for every $1 < p < \infty$.

This inequality will be proven in detail in a later section. Until then, we will use freely all these logarithmic bounds.

Banach estimates for $C_{2,d}^{l,r_1,r_2,\tilde{n},\tilde{n}_1,\tilde{n},\tilde{n}_1,\tilde{\alpha},\tilde{\beta},t_1,t_2}$. Given the logarithmic bounds for the shifted maximal and square functions described earlier, it is not difficult to see (as in [20]) that the operator $C_{2,d}^{l,r_1,r_2,\tilde{n},\tilde{n}_1,\tilde{n},\tilde{n}_1,\tilde{\alpha},\tilde{\beta},t_1,t_2}$ is indeed bounded from $L^{s_1} \times \dots \times L^{s_l}$ into L^s as long as $1/s_1 + \dots + 1/s_l = 1/s$ and $1 < s_1, \dots, s_l, s < \infty$ with operatorial bounds no greater than

$$\left(C < r_1 > < r_2 > \log < \tilde{n} > \log < \tilde{n} > \log < \tilde{n}_1 > \log < \tilde{n}_1 > \log < [t_1] > \log < [t_2] > \right)^{2l}. \quad (35)$$

And this contribution is perfect, given the extra factors $2^{r_1}, 2^{r_2}$ that appeared before (recall that both r_1, r_2 are negative in our case) and the quadratic decay in $\tilde{n}, \tilde{n}_1, \tilde{n}, \tilde{n}_1$.

⁵The logarithmical bounds for M_n are due to Stein and can be found in [24] Chapter II.

Quasi-Banach estimates for $C_{2,d}^{l,r_1,r_2,\widetilde{n},\widetilde{n}_1,\widetilde{n},\widetilde{n}_1,\vec{\alpha},\vec{\beta},t_1,t_2}$. Assume now that the index s above satisfies $0 < s < \infty$ and so it can be sub-unitary. We would like to estimate the boundedness constants of

$$C_{2,d}^{l,r_1,r_2,\widetilde{n},\widetilde{n}_1,\widetilde{n},\widetilde{n}_1,\vec{\alpha},\vec{\beta},t_1,t_2} : L^{s_1} \times \dots \times L^{s_l} \rightarrow L^s. \quad (36)$$

This time one has to discretize the operators in the x_1, x_2 variables and then take advantage of the general result in Theorem 2.1. Arguing as in [20] we see that the problem reduces to estimating expressions of type

$$\begin{aligned} & \frac{1}{2^{6r_1l}} \frac{1}{2^{6r_2l}} \sum_R \frac{1}{|R|^{(l-1)/2}} \left| \langle f, \Phi_{I_{[\widetilde{n}-t_1]}}^{1,0} \otimes \Phi_{J_{[\widetilde{n}-t_2]}}^{2,0} \rangle \right| \cdot \left| \langle f_1, \Phi_{I_{\widetilde{n}_1}}^{1,1} \otimes \Phi_{J_{\widetilde{n}_1}}^{2,1} \rangle \right| \\ & \prod_{j \in S} \left| \langle f_j, \Phi_{I_{[(\widetilde{n}-t_1)\alpha_j]}}^{1,j} \otimes \Phi_{J_{[(\widetilde{n}-t_2)\beta_j]}}^{2,j} \rangle \right| \cdot \left| \langle f_{d+1}, \Phi_I^{1,d+1} \otimes \Phi_J^{2,d+1} \rangle \right| \end{aligned}$$

where the sum runs over dyadic rectangles of the form $R = I \times J$. By applying Theorem 2.1 we see that the operatorial norms of (36) can be majorized by

$$\left(2^{-6r_1} 2^{-6r_2} \log < \widetilde{n} > \log < \widetilde{n} > \log < \widetilde{n}_1 > \log < \widetilde{n}_1 > \log < [t_1] > \log < [t_2] > \right)^{2l}$$

and the same is true for all its adjoint operators. In the end, by using the same interpolation argument as in [20], one can see that the operator $C_{2,d}^{l,r_1,r_2,\widetilde{n},\widetilde{n}_1,\widetilde{n},\widetilde{n}_1,\vec{\alpha},\vec{\beta},t_1,t_2}$ satisfies the inequality (9) with bounds that are clearly acceptable in (29) as desired.

These complete the discussion of Case $1'_a \otimes$ Case $1'_a$. The rest of the cases can be treated similarly after certain adjustments. Since all of these adjustments have been described carefully in [20], the only thing that is left is to realize that they work equally well in our *tensor product* framework. The straightforward (but quite delicate) details are left to the reader.

4. PROOF OF THEOREM 2.1

The proof of Theorem 2.1 is based on the method developed in [21] and [22]. First, we need to recall the following lemma whose detailed proof can be found in [22].

Lemma 4.1. *Let $J \subseteq \mathbf{R}$ be an arbitrary interval. Then, every bump function ϕ_J adapted to J can be written as*

$$\phi_J = \sum_{k \in \mathbf{N}} 2^{-1000lk} \phi_J^k \quad (37)$$

where for each $k \in \mathbf{N}$, ϕ_J^k is also a bump adapted to J but with the additional property that $\text{supp}(\phi_J^k) \subseteq 2^k J$ ⁶. Moreover, if we assume $\int_{\mathbf{R}} \phi_J(x) dx = 0$ then all the functions ϕ_J^k can be chosen so that $\int_{\mathbf{R}} \phi_J^k(x) dx = 0$ for every $k \in \mathbf{N}$.

⁶ $2^k J$ is the interval having the same center as J and whose length is $2^k |J|$.

Fix now the normalized functions f_1, \dots, f_l and the set E as in (13). Using the above lemma, one can estimate the $(l + 1)$ - linear form on the left hand side of (13) as

$$|\Lambda_{\mathcal{R}}(f_1, \dots, f_{l+1})| \leq \sum_{\vec{k} \in \mathbf{N}^2} 2^{-1000l|\vec{k}|} \sum_{R \in \mathcal{R}} \frac{1}{|R|^{(l-1)/2}} |\langle f_1, \Phi_{R_{n_1}}^1 \rangle| \dots |\langle f_l, \Phi_{R_{n_l}}^l \rangle| |\langle f_{l+1}, \Phi_R^{l+1, \vec{k}} \rangle| \quad (38)$$

where the new functions $\Phi_R^{l+1, \vec{k}}$ have basically the same structure as the old Φ_R^{l+1} but they also have the additional property that $\text{supp}(\Phi_R^{l+1, \vec{k}}) \subseteq 2^{\vec{k}}R$. We denoted by $2^{\vec{k}}R := 2^{k_1}I \times 2^{k_2}J$, $\vec{k} = (k_1, k_2)$ and $|\vec{k}| = k_1 + k_2$.

As before, the form (38) will be majorized later on by *tensorizing* two separate $l^2 \times l^2 \times l^\infty \times \dots \times l^\infty$ estimates with respect to parameters I and J . As a consequence, for every index $1 \leq j \leq l + 1$ there are *hybrid square and maximal functions* naturally attached to that position which we denote by $(M - S)_j$. More specifically $(M - S)_j$ can be either the *discrete variant* of $M_{n_j^1}M_{n_j^2}$ or $S_{n_j^1}S_{n_j^2}$ or $M_{n_j^1}S_{n_j^2}$ or $S_{n_j^1}M_{n_j^2}$ depending on the positions of the corresponding Ψ functions. For simplicity, we do not write explicitly the dependence of these functions $(M - S)_j$ on the *shifting* parameters n_j . Recall also that each of them comes with a boundedness constant which is no greater than $O(\log^2 \langle n_j^1 \rangle \log^2 \langle n_j^2 \rangle)^7$.

We construct now an exceptional set as follows. For each $\vec{k} \in \mathbf{N}^2$ define

$$\Omega_{-5|\vec{k}|} = \bigcup_{j=1}^l \{(x, y) \in \mathbf{R}^2 : (M - S)_j(f_j)(x, y) > C2^{5|\vec{k}|} \log^2 \langle n_j^1 \rangle \log^2 \langle n_j^2 \rangle\}. \quad (39)$$

Also, define

$$\tilde{\Omega}_{-5|\vec{k}|} = \{(x, y) \in \mathbf{R}^2 : MM(\chi_{\Omega_{-5|\vec{k}|}})(x, y) > \frac{1}{2l}\} \quad (40)$$

and then

$$\tilde{\tilde{\Omega}}_{-5|\vec{k}|} = \{(x, y) \in \mathbf{R}^2 : MM(\chi_{\tilde{\Omega}_{-5|\vec{k}|}})(x, y) > \frac{1}{2|\vec{k}|}\}. \quad (41)$$

Finally, we denote by

$$\Omega = \bigcup_{\vec{k} \in \mathbf{N}^2} \tilde{\tilde{\Omega}}_{-5|\vec{k}|}.$$

It is clear that $|\Omega| < 1/2$ if C is a big enough constant, which we fix from now on. Then, define $E' := E \setminus \Omega$ and observe that $|E'| \sim 1$.

Fix then $\vec{k} \in \mathbf{N}^2$ and look at the corresponding inner sum in (38). We split it into two parts as follows. Part I sums over those rectangles R with the property that

$$R \cap \tilde{\tilde{\Omega}}_{-5|\vec{k}|}^c \neq \emptyset \quad (42)$$

while Part II sums over those rectangles with the property that

⁷It is a *standard* fact that the continuous and the discrete variants of these operators behave similarly.

$$R \cap \tilde{\Omega}_{-5|\vec{k}|}^c = \emptyset. \quad (43)$$

We observe that Part II is identically equal to zero, because if $R \cap \tilde{\Omega}_{-5|\vec{k}|}^c \neq \emptyset$ then $R \subseteq \tilde{\Omega}_{-5|\vec{k}|}$ and in particular this implies that $2^{\vec{k}}R \subseteq \tilde{\Omega}_{-5|\vec{k}|}$ which is a set disjoint from E' . It is therefore enough to estimate Part I only.

Since $R \cap \tilde{\Omega}_{-5|\vec{k}|}^c \neq \emptyset$, it follows that $\frac{|R \cap \tilde{\Omega}_{-5|\vec{k}|}^c|}{|R|} \leq \frac{1}{2^l}$ or equivalently, $|R \cap \tilde{\Omega}_{-5|\vec{k}|}^c| > \frac{2^{l-1}}{2^l}|R|$.

We are now going to describe $l+1$ decomposition procedures, one for each function f_j for $1 \leq j \leq l+1$. Later on, we will combine them, in order to estimate our sum.

Independently, for every $1 \leq j \leq l$, define

$$\Omega_{-5|\vec{k}|+1}^j = \{(x, y) \in \mathbf{R}^2 : (M - S)_j(f_j)(x, y) > \frac{C2^{5|\vec{k}|} \log^2 \langle \mathbf{n}_j^1 \rangle \log^2 \langle \mathbf{n}_j^2 \rangle}{2^1}\}$$

and set

$$\mathcal{R}_{-5|\vec{k}|+1}^j = \{R \in \mathcal{R} : |R \cap \Omega_{-5|\vec{k}|+1}^j| > \frac{1}{2^l}|R|\},$$

then define

$$\Omega_{-5|\vec{k}|+2}^j = \{(x, y) \in \mathbf{R}^2 : (M - S)_j(f_j)(x, y) > \frac{C2^{5|\vec{k}|} \log^2 \langle \mathbf{n}_j^1 \rangle \log^2 \langle \mathbf{n}_j^2 \rangle}{2^2}\}$$

and set

$$\mathcal{R}_{-5|\vec{k}|+2}^j = \{R \in \mathcal{R} \setminus \mathcal{R}_{-5|\vec{k}|+1}^j : |R \cap \Omega_{-5|\vec{k}|+2}^j| > \frac{1}{2^l}|R|\},$$

and so on. The constant $C > 0$ is the one in the definition of the set E' above. Since there are finitely many rectangles, this algorithm ends after a while, producing the sets $\{\Omega_{s_j}^j\}$ and $\{\mathcal{R}_{s_j}^j\}$ such that $\mathcal{R} = \cup_{s_j} \mathcal{R}_{s_j}$.

We would clearly like to have such a decomposition available for the last function f_{l+1} as well. To do this, we first need to construct the analogue of the set $\Omega_{-5|\vec{k}|}$, for it. Pick $N > 0$ a big enough integer such that for every $R \in \mathcal{R}$ we have $|R \cap \Omega_{-N}^{l+1}| > \frac{2^{l-1}}{2^l}|R|$ where we defined

$$\Omega_{-N}^{l+1} = \{(x, y) \in \mathbf{R}^2 : (M - S)_{l+1}(f_{l+1})(x, y) > C2^N\}.$$

Then, similarly to the previous algorithms, we define

$$\Omega_{-N+1}^{l+1} = \{(x, y) \in \mathbf{R}^2 : (M - S)_{l+1}(f_{l+1})(x, y) > \frac{C2^N}{2^1}\}$$

and set

$$\mathcal{R}_{-N+1}^{l+1} = \{R \in \mathcal{R} : |R \cap \Omega_{-N+1}^{l+1}| > \frac{1}{2^l}|R|\},$$

then define

$$\Omega_{-N+2}^{l+1} = \{(x, y) \in \mathbf{R}^2 : (M - S)_{l+1}(f_{l+1})(x, y) > \frac{C2^N}{2^2}\}$$

and set

$$\mathcal{R}_{-N+2}^{l+1} = \{R \in \mathcal{R} \setminus \mathcal{R}_{-N+1}^{l+1} : |R \cap \Omega_{-N+2}^{l+1}| > \frac{1}{2l}|R|\},$$

and so on, constructing the sets $\{\Omega_{s_{l+1}}^{l+1}\}$ and $\{\mathcal{R}_{s_{l+1}}^{l+1}\}$ such that $\mathcal{R} = \cup_{s_{l+1}} \mathcal{R}_{s_{l+1}}^{l+1}$.

Then we write Part I as

$$\sum_{s_1, \dots, s_l > -5|\vec{k}|, s_{l+1} > -N} \sum_{R \in \mathcal{R}_{s_1, \dots, s_{l+1}}} \frac{1}{|R|^{(l+1)/2}} |\langle f_1, \Phi_{R_{n_1}}^1 \rangle| \dots |\langle f_l, \Phi_{R_{n_l}}^l \rangle| |\langle f_{l+1}, \Phi_R^{l+1, \vec{k}} \rangle| |R|, \quad (44)$$

where $\mathcal{R}_{s_1, \dots, s_{l+1}} := \mathcal{R}_{s_1}^1 \cap \dots \cap \mathcal{R}_{s_{l+1}}^{l+1}$. Now, if R belongs to $\mathcal{R}_{s_1, \dots, s_{l+1}}$ this means in particular that R has not been selected at either of the previous $s_j - 1$ steps for $1 \leq j \leq l+1$, which means that $|R \cap \Omega_{s_1-1}^1| \leq \frac{1}{2l}|R|$, ..., $|R \cap \Omega_{s_{l+1}-1}^{l+1}| \leq \frac{1}{2l}|R|$ or equivalently $|R \cap \Omega_{s_1-1}^{1c}| > \frac{2l-1}{2l}|R|$, ..., $|R \cap \Omega_{s_{l+1}-1}^{l+1c}| > \frac{2l-1}{2l}|R|$. But this implies that

$$|R \cap \Omega_{s_1-1}^{1c} \cap \dots \cap \Omega_{s_{l+1}-1}^{l+1c}| > \frac{1}{2}|R|. \quad (45)$$

In particular, using (45), the term in (44) is smaller than

$$\begin{aligned} & \sum_{s_1, \dots, s_l > -5|\vec{k}|, s_{l+1} > -N} \sum_{R \in \mathcal{R}_{s_1, \dots, s_{l+1}}} \frac{1}{|R|^{(l+1)/2}} |\langle f_1, \Phi_{R_{n_1}}^1 \rangle| \dots |\langle f_l, \Phi_{R_{n_l}}^l \rangle| |\langle f_{l+1}, \Phi_R^{l+1, \vec{k}} \rangle| |R \cap \Omega_{s_1-1}^{1c} \cap \dots \cap \Omega_{s_{l+1}-1}^{l+1c}| = \\ & \sum_{s_1, \dots, s_l > -5|\vec{k}|, s_{l+1} > -N} \int_{\Omega_{s_1-1}^{1c} \cap \dots \cap \Omega_{s_{l+1}-1}^{l+1c}} \sum_{R \in \mathcal{R}_{s_1, \dots, s_{l+1}}} \frac{1}{|R|^{(l+1)/2}} |\langle f_1, \Phi_{R_{n_1}}^1 \rangle| \dots |\langle f_l, \Phi_{R_{n_l}}^l \rangle| |\langle f_{l+1}, \Phi_R^{l+1, \vec{k}} \rangle| \chi_R(x, y) dx dy \\ & \lesssim \sum_{s_1, \dots, s_l > -5|\vec{k}|, s_{l+1} > -N} \int_{\Omega_{s_1-1}^{1c} \cap \dots \cap \Omega_{s_{l+1}-1}^{l+1c} \cap \Omega_{\mathcal{R}_{s_1, \dots, s_{l+1}}}} \prod_{j=1}^{l+1} (M-S)_j(f_j)(x, y) dx dy \\ & \lesssim \sum_{s_1, \dots, s_l > -5|\vec{k}|, s_{l+1} > -N} 2^{5l|\vec{k}|} \prod_{j=1}^l \log^2 \langle \mathbf{n}_j^1 \rangle \log^2 \langle \mathbf{n}_j^2 \rangle > 2^{-s_1} \cdot \dots \cdot 2^{-s_{l+1}} |\Omega_{\mathcal{R}_{s_1, \dots, s_{l+1}}}|, \end{aligned} \quad (46)$$

where

$$\Omega_{\mathcal{R}_{s_1, \dots, s_{l+1}}} := \bigcup_{R \in \mathcal{R}_{s_1, \dots, s_{l+1}}} R.$$

On the other hand we can write

$$\begin{aligned} |\Omega_{\mathcal{R}_{s_1, \dots, s_{l+1}}}| & \leq |\Omega_{\mathcal{R}_{s_1}}^1| \leq |\{(x, y) \in \mathbf{R}^2 : MM(\chi_{\Omega_{s_1}^1})(x, y) > \frac{1}{2l}\}| \\ & \lesssim |\Omega_{s_1}^1| = |\{(x, y) \in \mathbf{R}^2 : (M-S)_1(f_1)(x, y) > \frac{C(\vec{k}, \mathbf{n}_1)}{2^{s_1}}\}| \lesssim 2^{s_1 p_1}. \end{aligned}$$

Similarly, we have

$$|\Omega_{\mathcal{R}_{s_1, \dots, s_{l+1}}}| \lesssim 2^{s_j p_j}$$

for every $1 \leq j \leq l$ and also

$$|\Omega_{\mathcal{R}_{s_1, \dots, s_{l+1}}}| \lesssim 2^{s_{l+1}\alpha},$$

for every $\alpha > 1$. Here we used the fact that all the operators $(M - S)_j$ are bounded on L^s as long as $1 < s < \infty$ and also that $|E'| \sim 1$. In particular, it follows that

$$|\Omega_{\mathcal{R}_{s_1, \dots, s_{l+1}}}| \lesssim 2^{s_1 p_1 \theta_1} \cdot \dots \cdot 2^{s_l p_l \theta_l} 2^{s_{l+1} \alpha \theta_{l+1}} \quad (47)$$

for any $0 \leq \theta_1, \dots, \theta_{l+1} < 1$, such that $\theta_1 + \dots + \theta_{l+1} = 1$.

Now we split the sum in (46) into

$$\begin{aligned} & \sum_{s_1, \dots, s_l > -5|\vec{k}|, s_{l+1} > 0} 2^{5l|\vec{k}|} \prod_{j=1}^l \log^2 < \mathbf{n}_j^1 > \log^2 < \mathbf{n}_j^2 > 2^{-s_1} \cdot \dots \cdot 2^{-s_{l+1}} |\Omega_{\mathcal{R}_{s_1, \dots, s_{l+1}}}| \\ & + \sum_{s_1, \dots, s_l > -5|\vec{k}|, 0 > s_{l+1} > -N} 2^{5l|\vec{k}|} \prod_{j=1}^l \log^2 < \mathbf{n}_j^1 > \log^2 < \mathbf{n}_j^2 > 2^{-s_1} \cdot \dots \cdot 2^{-s_{l+1}} |\Omega_{\mathcal{R}_{s_1, \dots, s_{l+1}}}|. \end{aligned} \quad (48)$$

To estimate the terms in (48) we use the inequality (47) as follows. First, we choose $\theta_1, \dots, \theta_l$ small enough so that $1 - p_j \theta_j > 0$ for every $1 \leq j \leq l$. Because of this, θ_{l+1} can become quite close to 1. To estimate the first term in (48) we pick α very close to 1 so that $1 - \alpha \theta_{l+1} > 0$, while to estimate the second term we pick α large enough so that $1 - \alpha \theta_{l+1} < 0$.

With these choices, the sum in (48) is at most $O(2^{100l|\vec{k}|} \prod_{j=1}^l \log^2 < \mathbf{n}_j^1 > \log^2 < \mathbf{n}_j^2 >)$ and this makes the expression in (38) to be $O(\prod_{j=1}^l \log^2 < \mathbf{n}_j^1 > \log^2 < \mathbf{n}_j^2 >)$ as desired, after summing over $\vec{k} \in \mathbf{N}^2$.

This ends our proof.

5. LOGARITHMICAL BOUNDS FOR THE SHIFTED HYBRID MAXIMAL AND SQUARE FUNCTIONS

To complete the proof of the main theorem, we need to demonstrate the logarithmical bounds that have been used for the *shifted hybrid maximal and square functions*. As we mentioned before, the arguments of [19] and [21] show that they would follow from the following logarithmical bound for the vector valued Fefferman-Stein inequality.

Theorem 5.1. *Let $n \in \mathbf{Z}$ be a fixed integer and denote by M_n the shifted maximal operator associated to n . Then, one has*

$$\|(\sum_{j=1}^N |M_n f_j|^2)^{1/2}\|_p \leq C_p \log^2 < n > \cdot \|(\sum_{j=1}^N |f_j|^2)^{1/2}\|_p \quad (49)$$

for every N and any $1 < p < \infty$.

Proof The proof is a combination of the classical argument of Fefferman and Stein [24] with the new ideas from [19]. A nice description of the Fefferman-Stein inequality is in Workman [26] and we follow that presentation closely. There are three cases. Clearly, n is supposed to be large, otherwise there is nothing to prove. Assume also that n is positive, since the negative case is completely similar.

Case 1: $p = 2$. This case is very simple and it follows immediately from the theorem in [19] which says that M_n is bounded on L^2 (and in fact on any L^p) with an operatorial bound of type $O(\log < n >)$.

Case 2: $p > 2$. To understand this case one first needs to observe the following lemma.

Lemma 5.2. *The following inequality holds*

$$\alpha \cdot \int_{\{x: M_n f(x) > \alpha\}} |\Phi(x)| dx \lesssim \int_{\mathbf{R}} |f(x)| \mathcal{M}_n \Phi(x) dx \quad (50)$$

for every $\alpha > 0$ and measurable functions f and Φ , where \mathcal{M}_n has been defined to be

$$\mathcal{M}_n \Phi(x) := \sum_{k=0}^{\lfloor \log_2 n \rfloor} M_{-2^k} \Phi(x).$$

Proof To prove this lemma we need to recall a few facts from [19]. Denote by I_n maximal dyadic intervals selected with the property that

$$\frac{1}{|I_n|} \int_{I_n} |f(x)| dx > \alpha. \quad (51)$$

Clearly, they are all disjoint and their union is equal to $\{x : Mf(x) > \alpha\}$. Each I_n comes with $\lfloor \log_2 n \rfloor$ dyadic intervals of the same length *attached* to it, denoted by $I_n^1, \dots, I_n^{\lfloor \log_2 n \rfloor}$. More precisely, I_n^k lies 2^k steps of length $|I_n|$ to the left of I_n . It has been observed in [19] that

$$\{x : M_n f(x) > \alpha\} \subseteq \bigcup_{I_n} I_n \cup I_n^1 \cup \dots \cup I_n^{\lfloor \log_2 n \rfloor}.$$

Using these, one can majorize the left hand side of (50) by

$$\begin{aligned} & \alpha \cdot \sum_{I_n} \sum_{k=1}^{\lfloor \log_2 n \rfloor} \int_{I_n^k} |\Phi(x)| dx \\ & \lesssim \sum_{I_n} \sum_{k=1}^{\lfloor \log_2 n \rfloor} \left(\frac{1}{|I_n|} \int_{I_n} |f(y)| dy \right) \cdot \left(\int_{I_n^k} |\Phi(x)| dx \right) \\ & = \sum_{I_n} \sum_{k=1}^{\lfloor \log_2 n \rfloor} \left(\int_{I_n} |f(y)| dy \right) \cdot \left(\frac{1}{|I_n|} \int_{I_n^k} |\Phi(x)| dx \right). \end{aligned} \quad (52)$$

Now, for every $y \in I_n$ one can see that

$$\frac{1}{|I_n|} \int_{I_n^k} |\Phi(x)| dx \lesssim M_{-2^k} \Phi(y).$$

Using this in (52) one immediately obtains the desired (50). ■

The result of the above lemma implies that

$$M_n : L^1(\mathbf{R}, \mathcal{M}_n \Phi dx) \rightarrow L^{1,\infty}(\mathbf{R}, |\Phi| dx).$$

Since Φ can be assumed not to be identically equal to zero, we know that $\mathcal{M}_n \Phi > 0$. In particular, we also have the trivial bound

$$M_n : L^\infty(\mathbf{R}, \mathcal{M}_n \Phi dx) \rightarrow L^\infty(\mathbf{R}, |\Phi| dx)$$

and by interpolation, we obtain the following L^2 estimate

$$\int_{\mathbf{R}} |M_n f(x)|^2 |\Phi(x)| dx \lesssim \int_{\mathbf{R}} |f(x)|^2 \mathcal{M}_n \Phi(x) dx. \quad (53)$$

Coming back to the proof of Case 2, since $p > 2$ we know that $q := p/2 > 1$. By picking an appropriate $\|\Phi\|_{q'} = 1$ and relying on the previous (53), one can write

$$\begin{aligned} \|(\sum_{j=1}^N |M_n f_j|^2)^{1/2}\|_p^2 &= \|\sum_{j=1}^N |M_n f_j|^2\|_q = \int_{\mathbf{R}} (\sum_{j=1}^N |M_n f_j|^2) |\Phi| dx \\ &\lesssim \int_{\mathbf{R}} (\sum_{j=1}^N |f_j|^2) \mathcal{M}_n \Phi dx \lesssim \|\sum_{j=1}^N |f_j|^2\|_q \cdot \|\mathcal{M}_n \Phi\|_{q'} \lesssim \|(\sum_{j=1}^N |f_j|^2)^{1/2}\|_p^2 \cdot \|\mathcal{M}_n\|_{q' \rightarrow q'}. \end{aligned}$$

On the other hand, from the definition of \mathcal{M}_n and the result of [19], we know that

$$\|\mathcal{M}_n\|_{q' \rightarrow q'} \leq \sum_{k=1}^{\lfloor \log_2 n \rfloor} \|\mathcal{M}_{-2^k}\|_{q' \rightarrow q'} \lesssim \sum_{k=1}^{\lfloor \log_2 n \rfloor} \log 2^k \lesssim \log^2 < n >$$

which completes Case 2.

Case 3: $1 < p < 2$. The idea here is to prove the following end point case

$$\|(\sum_{j=1}^N |M_n f_j|^2)^{1/2}\|_{1,\infty} \lesssim \log^2 < n > \cdot \|(\sum_{j=1}^N |f_j|^2)^{1/2}\|_1 \quad (54)$$

directly and then to apply standard vector valued interpolation with the corresponding L^2 estimate.

To prove (54), let $\alpha > 0$ and denote by $F(x) := (\sum_{j=1}^N |f_j(x)|^2)^{1/2}$. Select maximal dyadic intervals I_n with the property

$$\frac{1}{|I_n|} \int_{I_n} F(x) dx > \alpha.$$

As before, we like to think of each I_n as being related to the dyadic interval I , having the same length as I_n and lying n steps of length $|I_n|$ to the left of it. If we denote by $\Omega := \bigcup_I I_n$ one has as usual

$$|\Omega| = \sum_I |I_n| \leq \frac{1}{\alpha} \sum_I \int_{I_n} F(x) dx \leq \frac{1}{\alpha} \|F\|_1. \quad (55)$$

Observe that $F \leq \alpha$ on Ω^c and also that

$$\alpha < \frac{1}{|I_n|} \int_{I_n} F(x) dx \leq 2\alpha$$

because of the maximality of I_n .

Split now each f_k as $f_k = f'_k + f''_k$ where $f'_k := f_k \chi_{\Omega^c}$ and $f''_k := f_k \chi_{\Omega}$.

Contribution of $\{f'_k\}$. One can write

$$\begin{aligned} |\{x : (\sum_j |M_n f'_j(x)|^2)^{1/2} > \alpha/2\}| &\leq \frac{1}{\alpha^2} \|(\sum_j |M_n f'_j(x)|^2)^{1/2}\|_2^2 \\ &\lesssim \frac{1}{\alpha^2} \log^2 < n > \|(\sum_j |f'_j(x)|^2)^{1/2}\|_2^2 \lesssim \log^2 < n > \frac{1}{\alpha^2} \int_{\Omega^c} F^2(x) dx \leq \log^2 < n > \frac{1}{\alpha} \|F\|_1 \end{aligned}$$

as desired.

Contribution of $\{f''_k\}$. To estimate the corresponding contribution for $\{f''_k\}$, we have to be a bit more careful. Define first the functions g_k by

$$g_k := \sum_I \left(\frac{1}{|I_n|} \int_{I_n} |f_k(x)| dx \right) \cdot \chi_{I_n} \quad (56)$$

and after that $G(x) := (\sum_j |g_j(x)|^2)^{1/2}$. Fix $x \in I_n$ and observe that by Minkowski's inequality one can write

$$G(x) = \left(\sum_j \left(\frac{1}{|I_n|} \int_{I_n} |f_j(y)| dy \right)^2 \right)^{1/2} \leq \frac{1}{|I_n|} \int_{I_n} \left(\sum_j |f_j(y)|^2 \right)^{1/2} dy = \frac{1}{|I_n|} \int_{I_n} F(y) dy \leq 2\alpha.$$

Using that G is supported in Ω and arguing as before, we have

$$\begin{aligned} |\{x : (\sum_j |M_n g_j(x)|^2)^{1/2} > \alpha/2\}| &\lesssim \frac{1}{\alpha^2} \log^2 < n > \|(\sum_j |g_j(x)|^2)^{1/2}\|_2^2 \\ &= \frac{1}{\alpha^2} \log^2 < n > \|G\|_2^2 \lesssim \log^2 < n > |\Omega| \leq \frac{1}{\alpha} \log^2 < n > \|F\|_1. \end{aligned}$$

We would like now to compare $M_n f''_k(x)$ with $M_n g_k(x)$ if possible. Denote by $\widetilde{\Omega}$ the set

$$\widetilde{\Omega} := \bigcup_I 3I_n \cup 3I_n^1 \cup \dots \cup 3I_n^{[\log_2 n]}$$

and observe that $|\widetilde{\Omega}| \lesssim \log < n > |\Omega|$. We will prove that for every $x \in \widetilde{\Omega}^c$ one has

$$M_n f''_k(x) \leq M_n g_k(x) \quad (57)$$

and this will clearly allow us to reduce the contribution of $\{f''_k\}$ to the contribution of $\{g_k\}$ which has been understood earlier. Fix $x \in \widetilde{\Omega}^c$ and $x \in J$ a dyadic interval, so that the corresponding $\frac{1}{|J_n|} \int_{J_n} |f''_k(y)| dy$ is different from zero. In particular, J_n has to intersect Ω which is the support

of f_k'' . Suppose now that I is so that $J_n \cap I_n \neq \emptyset$. Then, one must have $I_n \subseteq J_n$ as the other alternative $J_n \subseteq I_n$ is not possible since $x \in \widetilde{\Omega}^c$. But this implies that

$$\frac{1}{|J_n|} \int_{J_n} |f_k''(x)| dx = \frac{1}{|J_n|} \int_{J_n} |g_k(x)| dx$$

which is enough to guarantee (57) and end our proof. ■

6. GENERALIZATIONS

The goal of this section is to point out that virtually all the earlier generalizations that we described in the previous [19] and [20], have *natural* extensions in this *multi-parameter world* and can be proved by the same method. We give here just two *samples* and leave the rest (and the straightforward details) to the imaginative reader. Suppose for simplicity that we are in \mathbf{R}^2 and denote by $D_1 := \frac{\partial}{\partial x_1}$ and $D_2 := \frac{\partial}{\partial x_2}$. A direct computation shows that the *double commutator* $[[D_2], [[D_1], A]]$ can be rewritten as

$$[[D_2], [[D_1], A]]f(x) = p.v. \int_{\mathbf{R}^2} f(x+t) \left(\frac{\Delta_{t_1}^{(1)}}{t_1} \circ \frac{\Delta_{t_2}^{(2)}}{t_2} A(x) \right) \frac{dt_1}{t_1} \frac{dt_2}{t_2} \quad (58)$$

which is precisely the *bidisc* extension of the first commutator of Calderón. There is of course a similar formula available in every dimension.

Theorem 6.1. *Let a_1, \dots, a_n be real numbers, all different from zero. The expression*

$$p.v. \int_{\mathbf{R}^n} f(x+t) \left(\frac{\Delta_{a_1 t_1}^{(1)}}{t_1} \circ \dots \circ \frac{\Delta_{a_n t_n}^{(n)}}{t_n} A(x) \right) \frac{dt_1}{t_1} \dots \frac{dt_n}{t_n}$$

viewed as a bilinear map in f and $\frac{\partial^n A}{\partial x_1 \dots \partial x_n}$ is bounded from $L^p \times L^q$ into L^r for every $1 < p, q \leq \infty$ with $1/p + 1/q = 1/r$ and $1/2 < r < \infty$.

The particular case $q = \infty$ is in Journé [10] but the rest of the estimates seem to be new.

Then, one can also observe by a direct calculation that

$$\begin{aligned} & [[D_2], [[D_1], [[D_2], [[D_1], A]]]]f(x) \\ &= p.v. \int_{\mathbf{R}^4} f(x+t+s) \left(\frac{\Delta_{t_1}^{(1)}}{t_1} \circ \frac{\Delta_{t_2}^{(2)}}{t_2} \circ \frac{\Delta_{s_1}^{(1)}}{s_1} \circ \frac{\Delta_{s_2}^{(2)}}{s_2} A(x) \right) \frac{dt_1}{t_1} \frac{dt_2}{t_2} \frac{ds_1}{s_1} \frac{ds_2}{s_2} \end{aligned}$$

which is the *bidisc* analogue of an operator introduced in [20]. As we promised, we record now the following theorem.

Theorem 6.2. *Let F be an analytic function on a disc of a certain radius centered at the origin in the complex plane and A a complex valued function in \mathbf{R}^2 so that $\frac{\partial^4 A}{\partial x_1^2 \partial x_2^2} \in L^\infty(\mathbf{R}^2)$ with an L^∞ norm strictly smaller than the radius of convergence of F . Then, the linear operator*

$$f \rightarrow p.v. \int_{\mathbf{R}^4} f(x+t+s) F \left(\frac{\Delta_{t_1}^{(1)}}{t_1} \circ \frac{\Delta_{t_2}^{(2)}}{t_2} \circ \frac{\Delta_{s_1}^{(1)}}{s_1} \circ \frac{\Delta_{s_2}^{(2)}}{s_2} A(x) \right) \frac{dt_1}{t_1} \frac{dt_2}{t_2} \frac{ds_1}{s_1} \frac{ds_2}{s_2}$$

is bounded on $L^p(\mathbb{R}^2)$ for every $1 < p < \infty$.

REFERENCES

- [1] Calderón A., *Commutators of singular integral operators*, Proc. Nat. Acad. Sci. U.S.A., vol. 53, 1092-1099, [1965].
- [2] Calderón A., *Commutators, singular integrals on Lipschitz curves and applications*, Proc. Int. Congress of Math., Helsinki, 1978. Academia Scientiarum Fennica, Helsinki, 85-96, [1980].
- [3] Christ M., Journé J.L., *Polynomial growth estimates for multilinear singular integral operators*, Acta Math., vol. 159, 51-80, [1987].
- [4] Cohen J., Gosselin J., *A BMO estimate for multilinear singular integrals*, Illinois J. Math., vol. 30, 445-465, [1986].
- [5] Coifman R. R., Jones P., Semmes S., *Two elementary proofs of the L^2 -boundedness of the Cauchy integral on Lipschitz curves*, J. Amer. Math. Soc., vol. 2, 553-564, [1989].
- [6] Coifman R. R., Meyer, Y., *Wavelets. Calderón-Zygmund and multilinear operators*, Cambridge Studies in advanced mathematics, [1997].
- [7] Coifman R., McIntosh, A. and Meyer, Y., *L'intégral de Cauchy définit un opérateur borné sur les courbes Lipschitziennes*, Ann of Math., vol. 116, 361-387, [1982].
- [8] David G., Journé J.L., *A boundedness criterion for generalized Calderón-Zygmund operators*, Ann. of Math., vol. 120, 371-397, [1984].
- [9] David G., Journé J.L., Semmes S., *Opérateurs Calderón-Zygmund, fonctions para-accréitive et interpolation*, Revista Mat. Iberoamericana, vol. 1, 1-56, [1985].
- [10] Journé J.L., *Calderón-Zygmund operators on product spaces*, Revista Mat. Iberoamericana, vol. 1, 55-91, [1985].
- [11] Journé J.L., *Two problems of Calderón-Zygmund theory on product spaces*, Ann. de L'Institut Fourier, vol. 38, 111-132, [1988].
- [12] Fefferman C., *Recent progresss in classical Fourier analysis*, Proceedings of the ICM, Vancouver, 1974, vol. 1, 95-118, Canad. Math. Congress, Montreal, QE, [1975].
- [13] Kenig C., *Weighted H^p spaces on Lipschitz domains*, Amer. J. Math., vol. 102, 129-163, [1980].
- [14] Lacey M., Thiele C., *L^p estimates for the bilinear Hilbert transform for $2 < p < \infty$* , Ann. of Math., vol. 146, 693-724, [1997].
- [15] Lacey M., Thiele C., *On Calderón's conjecture*, Ann. of Math., vol. 149, 475-496, [1999].
- [16] Mitrea M., *Clifford Wavelets, Singular integrals and Hardy spaces*, LNM, vol. 1575, Springer Verlag, [1994].
- [17] Murai T., *A real variable method for the Cauchy transform and analytic capacity*, L.N.M. vol 1307, Springer-Verlag, Berlin, [1984].
- [18] Muscalu C., *Paraproducts with flag singularities I. A case study*, Revista Mat. Iberoamericana, vol. 23, 705-742, [2007].
- [19] Muscalu C., *Calderón commutators and the Cauchy integral on Lipschitz curves revisited I. First commutator and generalizations*, Preprint.
- [20] Muscalu C., *Calderón commutators and the Cauchy integral on Lipschitz curves revisited II. Cauchy integral and generalizations*, Preprint.
- [21] Muscalu C., Pipher J., Tao T., Thiele C., *Bi-parameter paraproducts*, Acta Math., vol. 193, 269-296, [2004].
- [22] Muscalu C., Pipher J., Tao T., Thiele C., *Multi-parameter paraproducts*, Revista Mat. Iberoamericana, vol. 22, 963-976, [2006].
- [23] Muscalu C., Tao T., Thiele C., *Multi-linear operators given by singular multipliers*, J. Amer. Math. Soc., vol. 15, pages 469-496, [2001].
- [24] Stein E., *Harmonic analysis: real-variable methods, orthogonality and oscillatory integrals*, Princeton University Press, [1993].
- [25] Verdera J., *L^2 boundedness of the Cauchy integral and Menger curvature*, Contemporary Mathematics, vol. 277, 139-158, [2001].
- [26] Workman, J., *End-Point estimates and multi-parameter paraproducts on higher dimensional tori*, Ph.D. Thesis, Cornell University, [2008].

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NY 14853
E-mail address: `camil@math.cornell.edu`